PB84-168426

Neighborhood Complexities and Symmetry of Chemical Graphs and Their Biological Applications

Minnesota Univ.-Duluth. Dept. of Chemistry

Prepared for

Environmental Research Lab.-Duluth, MN

1984

U.S. Department of Commerce
National Technical Information Service
NTIS

PB84-168426

EPA-600/D-84-065
1984

NEIGHBORHOOD COMPLEXITIES AND SYMMETRY OF CHEMICAL GRAPHS AND THEIR BIOLOGICAL APPLICATIONS

by

A. B. Roy, S. C. Basak, D. K. Harriss, and V. R. Magnuson
University of Minnesota, Duluth
Duluth, MN 55812

CR 807566

EPA Project Officer
G. D. Veith

ENVIRONMENTAL RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
DULUTH, MN 55804

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
EPA-600/D-84-065		PB84 168426

4. TITLE AND SUBTITLE	5. REPORT DATE
Neighborhood Complexities and Symmetry of Chemical Graphs and their Biological Applications	1984
	6. PERFORMING ORGAN ZATION CODE
7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
A.B. Roy, S.C. Basak, D.K. Harriss, and V.R. Magnuson	
9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT NO.
Department of Chemistry University of Minnesota-Duluth Duluth, Minnesota 55812	11. CONTRACT/GRANT NO. CR 807566
12. SPONSORING AGENCY NAME AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
U.S. Environmental Protection Agency Environmental Research Laboratory-Duluth 6201 Congdon Boulevard Duluth, Minnesota 55804	14. SPONSORING AGENCY CODE EPA/600/03

15. SUPPLEMENTARY NOTES

16. ABSTRACT

Quantitative measures of molecular complexity are calculated through the application of information-theoretic formalism on chemical graphs. The vertex set of a chemical graph is partitioned into disjoint subsets on the basis of the equivalence of various orders of closed neighborhoods and the information indices (IC_ν, SIC_ν, CIC_ν, and R_ν) are calculated. The applications of these indices in structure-activity correlations are discussed.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
RELEASE TO PUBLIC	UNCLASSIFIED	9
	20. SECURITY CLASS *(This page)*	22. PRICE
	UNCLASSIFIED	

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

NEIGHBORHOOD COMPLEXITIES AND SYMMETRY OF CHEMICAL GRAPHS AND THEIR BIOLOGICAL APPLICATIONS

A. B. Roy, S. C. Basak, D. K. Harriss, and V. R. Magnuson

Department of Chemistry, University of Minnesota, Duluth, Duluth Minnesota 55812, USA

Abstract. Quantitative measures of molecular complexity are calculated through the application of information-theoretic formalism on chemical graphs. The vertex set of a chemical graph is partitioned into disjoint subsets on the basis of the equivalence of various orders of closed neighborhoods and the information indices (IC_v, SIC_v, CIC_v, and R_v) are calculated. The applications of these indices in structure-activity correlations are discussed.

Keywords. Graph theory; symmetry; information-theoretic topological indices; modelling; probability.

INTRODUCTION

In recent years graph invariants have been found to be useful in chemical documentation, isomer discrimination, structure-property correlations and chemical structure-biological activity relationships (Balaban and co-workers, 1980; Kier and Hall, 1976; Randić and Wilkins, 1979). A numerical characteristic derived from a graph is usually called a graph invariant if for isomorphic graphs it remains an invariant. Different types of graph invariants, which condense the connectivity aspect, are available in the literature. The forms of such invariants may be a matrix, a polynomial, a sequence of numbers or a numerical index. For example, the characteristic polynomial of the adjacency matrix, the spectrum of a graph and the total number of different possible distances between pairs of vertices of a graph (called Wiener number) are all graph invariants. There are also various kinds of numerical indices associated with a graph which are usually called information-theoretic indices (Balaban and co-workers, 1980; Basak, Roy and Ghosh, 1980; Bonchev, Mekenyan and Trinajstić, 1981; Kier, 1980; Rashevsky, 1955; Sarkar, Roy and Sarkar, 1978). The invariants in the form of single numerical descriptors are found to be useful in chemical documentation and also as the criteria for the testing of isomorphism of graphs. Although many graph invariants have been reported in the literature, there does not exist a set of invariants which can recognize the isomorphism of arbitrary graphs (Read and Cornell, 1977). Yet different graph invariants have been successfully utilized in solving problems of chemistry and biology. Various graph invariants in the form of information-theoretic indices are being proposed in the recent literature. To derive these indices, a probability scheme is defined on the appropriate set of elements derived from the graph and the complexity (information content) of the discrete system is calculated using Shannon's (1949) formula.

In the present study, we propose a few information-theoretic graph invariants based on the concept of neighborhood of the chemical graph in which the multiplicity of the edges and the nature of the atoms are considered. These indices are called topological information indices because they are derived on the basis of different neighborhood-preserving groups for the chemical graph. The set of indices reported here arise out of a general scheme based on the earlier studies in this direction (Basak, Roy and Ghosh, 1980; Basak and co-workers, 1982; Basak, Harriss and Magnuson, 1983; Sarkar, Roy and Sarkar, 1978).

PRELIMINARIES OF GRAPHS AND THEIR INFORMATION CONTENT

By a graph $G = (V,E)$ we mean a finite set V (whose elements are called vertices) together with a set E of two element subsets of V (the elements of E are called edges). The set of vertices and edges are usually denoted by V(G) and E(G) respectively. Two vertices are called adjacent if they are connected by an edge. A multigraph is a graph that may contain more than one edge between a given pair of adjacent vertices; we shall use the term graph for a finite, undirected graph without self loops, but which may include multiple edges. A walk of length k in a graph G from vertex u to

u_{k+1} is a sequence of vertices $S = u_1, u_2, \ldots, u_k, u_{k+1}$ for which the edge $(a_i, a_{i+1}) \in E(G)$ for $i = 1, 2\text{--}k$; the walk is open or closed accordingly as $a_1 \neq a_{k+1}$ or $a_1 = a_{k+1}$ respectively. A path is an open walk in which all vertices are distinct. A graph is said to be connected if every pair of vertices are connected by a path. The distance $\partial(u,v)$ between vertices u, v in G is the length of any shortest path from u to v. The number of edges incident with a vertex u in a graph is called the degree (valency) of u and is denoted by deg u. For a given vertex u, the integer $e(u) = \max_{v \in V} \partial(u,v)$ measures the distance from u to the vertex (or vertices) most remote from u. e(u) is called eccentricity of u, and it is said that a vertex is relatively central if e(u) is relatively small. The radius, ρ, of G is defined as $\rho = \min_{u \in V} e(u) = \min_{u,v \in V} \max_{u,v \in V} \partial(u,v)$. If $e(u) = \rho$, u is called a center of G. The diameter of G is defined as

$$D(G) = \max_{u \in V} e(u) = \max_{u,v \in V} \partial(u,v)$$

Two graphs G_1 and G_2 are said to be isomorphic if there exists a one-one mapping ϕ of $V(G_1)$ onto $V(G_2)$ such that $(u,v) \in E(G_1)$ if and only if:

$$(u,v)\phi = (u\phi, v\phi) \in E(G_2)$$

Such a mapping ϕ is called an isomorphism of G_1 onto G_2. An automorphism of a graph G is a one-one mapping of V(G) onto itself which preserves the number of links joining any two. The set of all automorphisms of G forms a group. The automorphism group of a graph will be denoted here by Aut G. If Aut G is the group of a graph G with vertex set $V(G) = \{u_1, u_2, \ldots u_n\}$, then for each $\phi \in$ Aut G $u_i = u_j = u_{i\phi}$ for $1 \leq i \leq n$. Hence to each $\phi \in$ Aut G there corresponds a unique permutation of the elements 1,2--n. So Aut G is a subgroup of the symmetric group S_n of degree n. If H is a subgroup of S_n and $1 \leq i \leq n$, then $\{i\phi \mid \phi \in H\}$ is called an orbit or transitive set of H. The distinct orbits of H decompose the set of elements {1,2--n}. Therefore, Aut G, which is a subgroup of S_n where n is the number of vertices, partitions V(G) into disjoint classes.

A molecule can be modelled as a linear graph without self loop and multiple edges in which atoms represent vertices and covalent bonds as edges. This graph is usually called the topological graph of the molecule. This graphical representation of a molecule does not take into account different kinds of atoms or the multiplicity of edges. For quantum chemical calculations, the above-mentioned graph (in which hydrogen atoms are deleted) is found to be very useful; but, for other applications, e.g. chemical reactions, calculations of information content of molecular structure, the graph of a molecule should be refined. To do so we define the chemical graph of a molecule in the following manner: A chemical graph associated with a molecule is denoted by G(V,E) where $V = \{v_1, v_2 \text{--} v_m\}$ denotes the collection of atoms of the form $v_i = (a_i, d_i)$, where a_i stands for an alphanumeric symbol representing the chemical identity of v_i and d_i (a positive integer) denotes the valency of a_i. Chemical graph G(V,E) may be linear or multigraph without a self loop.

The measurement of complexity or redundancy of a chemical graph is an interesting problem as it quantitates a qualitative concept like symmetry. One of the approaches to such measurement, in which orbits of the topologically equivalent atoms are realized from the point group of the chemical structure, is information-theoretic in nature. Rashevsky (1955) first proposed such an information-theoretic measure, with the help of Shannon's formula, for measuring the information content of a molecule. Trucco (1956) subsequently showed that topologically equivalent vertices of a graph are the orbits of the associated automorphism group of the graph. Trucco defined the information content (topological) in the following way: Let G be a graph with $|V(G)| = n$ and let $A_1, A_2 \text{--} A_h$ be the orbits of Aut G. A finite probability distribution is then defined as

$$\begin{matrix} A_1 & A_2 & - & - & A_h \\ p_1 & p_2 & - & - & p_h \end{matrix}$$

where

$$\sum_{i=1}^{h} p_i = 1, \quad p_i \geq 0, \quad i = 1, 2\text{--}h; \quad p_i = \frac{|A_i|\phi}{n}, \quad 1 \leq i \leq h$$

The structural information content (topological information content), I_{top}, is defined as

$$I_{top} = -\sum_{i=1}^{h} p_i \log_2 p_i \text{ bits} \qquad (1)$$

This is the mean information content per vertex of the graph and the total information content is

$$n \times I_{top} = n \log_2 n - \sum_{i=1}^{h} n_i \log_2 n_i$$

Mowshowitz (1968a), however, pointed out that the index of relative complexity as defined by (1) is not a selective measure of information as it is understood in information theory. He proposed that I_{top} measures information content relative to a system of transformations under which the graph remains invariant. It would be more appropriate to treat (1) as a measure of non-homogeneity of distribution of vertices where p_i stands for the probability of finding an arbitrary chosen vertex of the graph in the i^{th} orbit. It is clear that the determination of equivalent vertices and associated information content of a graph depends on the criterion of equivalency, and so it

can be defined in many ways. Therefore, the information content of a graph is not unique. For a chemical graph, various decompositions of vertices (edges) are proposed by Bonchev and co-workers (1981) who showed their applications in isomer discrimination. For the characterization of branching of an alkane's carbon skeleton, I_{top} is found to be unsuitable. In all such calculations of different information theoretic topological indices, the hydrogen depleted graphs with carbon skeletons were taken to be models of a chemical structure. To take into account the chemical structures with more than one kind of atom and also edge multiplicity, we have redefined the chemical graph. In the present study, we give a general formalism for identifying equivalent vertices and associated information content relative to different possible groups from neighborhood consideration (which are subgroups of corresponding symmetric group, S_n). It is found that the information content of a chemical structure relative to the degree preserving group is I_{cc} (Bonchev, Kamenski and Kamenska, 1976) and I_{orb} or I_{top} is information content relative to the permutation group formed by those permutations which preserve not only the degree but also the highest possible neighbors.

GROUPS OF A CHEMICAL GRAPH AND RELATED INFORMATION CONTENT

Obviously G is undirected graph without self loop. Let Aut G be the automorphism group of a chemical graph G. It is clear from the definition that if $\alpha\varepsilon$Aut G and $v\varepsilon V(G)$, then deg $(v\alpha)$ = deg v. Hence Aut G is a subgroup of the degree preserving group $P_0(G)$. $P_0(G)$ consists of all permutations of V(G) such that for each $v\varepsilon V(G)$ and $\alpha\varepsilon P_0(G)$, deg v = deg $(v\alpha)$. The first neighborhood $\Gamma^1(v)$ of a vertex $v\varepsilon V(G)$ can be defined as follows:

$$\Gamma^1(v) = \{u \mid \partial(u,v) = 1\}$$

and the first order closed neighborhood $N^1(v)$ of v can be defined as

$$N^1(v) = (v)U\Gamma^1(v) = \Gamma^0(v)U\Gamma^1(v)$$

if (v) is taken as $\Gamma^0(v)$. Defining open sphere S(v,r) of v as $S(v,r) = \{u \mid \partial(u,v)<r\}$ where r is a positive integer and $S(v,o) = \phi$, it is shown that all such open spheres of the vertices of G form a neighborhood system of completely normal topological space where $0\leq r\leq 2$. Similar conclusions can be drawn if $0\leq r\leq \rho$ where ρ is the radius of the graph. Let us consider the set of all those permutations of $P_0(G)$ which preserve the first order neighbors in the sense that for all $\alpha\varepsilon P_0(G)$:

i) $|\Gamma^1(u)| = |\Gamma^1(u\alpha)|$

ii) For each vertex $v\varepsilon\Gamma_1^1(u)$ there exists a corresponding vertex $\omega\varepsilon\Gamma^1(u\alpha)$ such that the number of edges joining u and v is the same as the number of edges joining $u\alpha$ and ω.

iii) deg v = deg ω for all $u\varepsilon\Gamma^1(u)$ and $\omega\varepsilon\Gamma^1(u)$.

Evidently all such permutations form a subgroup $P_1(G)$ of the degree preserving group $P_0(G)$, i.e. $P_1(G)\subset P_0(G)$. In the similar manner, one can easily establish that there are ρ possible subgroups $P_0(G)$, $P_1(G)$--$P_\rho(G)$ of G such that $P_0(G) \subset P_1(G) \subset P_2(G)$--$\subset P_\rho(G)$. Since the group $P_\rho(G)$ contains all those permutations which preserve all successive first order neighbors in the above sense up to ρ, this group is identical with Aut (G). Hence we have the following theorem:

Th. 1 If $P_0(G)$, $P_1(G)$,--$P_\rho(G)$ are all different possible neighborhood preserving groups of an undirected multigraph, then

$$P_0(G) \subset P_1(G)\text{--}\subset P_\rho(G) = \text{Aut } G$$

It is clear that if $P_i(G) = P_{i+1}(G)$ for $0\leq i\leq\rho$, then $P_{i+1}(G) = P_{i+2}(G) = \ldots\ldots = P_\rho(G) = \text{Aut G}$.

To calculate the information content of G relative to these neighborhood preserving groups, we require the knowledge of the number and cardinalities of each such group. The number of orbits can be determined by Burnside's lemma (1911). The lemma states that the number of equivalent classes into which S is divided by the equivalence relation induced by a permutation group of symmetric group S_n is given by

$$\frac{1}{|G|}\sum_{\pi(G)}\psi(\pi)$$

where $\psi(\pi)$ = number of elements of S that are invariant under π. Determination of all such permutation groups is a difficult problem when the number of vertices becomes large. Here we propose a method to identify the equivalent vertices of a graph. Two vertices u_o and v_o of a chemical graph G are said to be equivalent with respect to the k^{th} order neighborhood if and only if:

i) deg u_o = deg v_o, u_o and v_o being atoms of the same chemical element.

ii) $|N^i(u_o)| = |N^i(v_o)|$ for i=1,2,...,k; $|N^i(u_o)|$ and $|N^i(v_o)|$ represent the cardinalities of $N^i(u_o)$ and $N^i(v_o)$ respectively;

iii) corresponding to each path $pu_o = u_o u_1 \ldots u_\ell$, $u_\ell\varepsilon\Gamma^\ell(u_o)$ for ℓ=1,2,...,k, there is a distinct path $pv_o = v_o v_1 \ldots v_\ell$, $v_\ell\varepsilon\Gamma^\ell(v_o)$ for ℓ=1,2,...,k, such that: (a) deg u_i=deg v_i, u_i and v_i being atoms of the same type for each i satisfying the relation i=1,2,...,k; and (b) $E(U_i,u_{i-1}) = E(v_i,v_{i-1})$ for i=1,2,..., k, where $E(u_i,u_{i-1})$ is the number of edges joining u_i and u_{i-1} and $E(v_i,v_{i-1})$ represents the number of edges connecting v_i and v_{i-1}.

This equivalence relation determines the equivalent vertices and determines the information content IC_ν relative to the group P_ν where $0<\nu<\rho$. It is easy to see that at least one of the orbits of P_ν will be a union of orbits of $P_{\nu+1}$ if $P_\rho \subset P_{\rho+1}$. Hence, by using Lemma 2.5 of Mowshowitz (1968b), the following theorem can be stated:

Th. 2 For an undirected multigraph G,

$$IC_0 \leq IC_1 \leq IC_2 \cdots \leq IC_\rho = I_{top}$$

where IC_ν denotes the topological information content of G relative to the neighborhood preserving group P_ν. It is easy to see that IC_0 is the same as I_{cc} of a chemical graph G and $IC_\rho = I_{orbit}$ of Rashevsky (1955).

Other topological indices can be easily derived as generalizations of SIC, CIC and RNSIC (Basak, Roy and Ghosh, 1980; Raychaudhury and co-workers, 1982):

i) The structural information content (SIC_ν) of G relative to permutation group P_ν of G is

$$SIC_\nu = \frac{IC_\nu}{\log_2 n} \quad \text{where } |V(G)| = n.$$

This may be looked at as a normalized IC_ν with the maximum information content $\log_2 n$ as normalizing factor.

ii) Taking into account the deviation of IC_ν from the maximum information content $\log_2 n$, one gets a new index

$$CIC_\nu = \log_2 n - IC_\nu$$

where CIC_ν stands for complementary information content of G.

iii) Another index can be derived from CIC_ν by normalizing using $\log_2 n$. This index measures the relative redundancy (R_ν) of G.

$$R_\nu = \frac{CIC_\nu}{\log_2 n} = 1 - \frac{IC_\nu}{\log_2 n} = 1 - SIC_\nu$$

R_ν is the same as the RNSIC index proposed by Basak, Roy and Ghosh (1980) when $\nu=1$.

SAMPLE CALCULATION OF INFORMATION - THEORETIC INDICES FROM A CHEMICAL GRAPH

To clarify the vertex partitioning into equivalent classes on the basis of neighborhood symmetry as described in the previous section and to illustrate the calculation of information-theoretic indices at various orders, 2-butenol will be used as an example. Figure 1 shows the vertex labeling for 2-butenol as well as the closed neighborhoods for the various orders, zero through three. Equivalent vertices, the probability, and IC_ν, SIC_ν, and CIC_ν for ν=0-3 are given in Table 1.

TABLE 1 Equivalent Vertices, Probability and Topological Indices for Order 0 through 3 for 2-Butenol

Order	Equivalent Vertices	Probability	Calculated Values
0	(C_1,C_2,C_3,C_4)	4/13	IC_0 =1.2389
	(O)	1/13	SIC_0=0.3348
	$(H_1,H_2,H_3,H_4,H_5,H_6,H_7,H_8)$	8/13	CIC_0=2.4615
1	(C_1)	1/13	
	(C_2,C_3)	2/13	
	(C_4)	1/13	IC_1 =2.0349
	(O)	1/13	SIC_1=0.5499
	$(H_1,H_2,H_3,H_4,H_5,H_6,H_7)$	7/13	CIC_1=1.6655
	(H_8)	1/13	
2	(C_1)	1/13	
	(C_2)	1/13	
	(C_3)	1/13	IC_2 =3.0270
	(C_4)	1/13	SIC_2=0.8180
	(O)	1/13	CIC_2=0.6735
	(H_1,H_2,H_3)	3/13	
	(H_4,H_5)	2/13	
	(H_6,H_7)	2/13	
	(H_8)	1/13	
3	(C_1)	1/13	
	(C_2)	1/13	
	(C_3)	1/13	
	(C_4)	1/13	IC_3 =3.1808
	(O)	1/13	SIC_3=0.8596
	(H_1,H_2,H_3)	3/13	CIC_3=0.5196
	(H_4)	1/13	
	(H_5)	1/13	
	(H_6,H_7)	2/13	
	(H_8)	1/13	

Labeling as in Fig. 1

If a vertex partitions into a class by itself at any order n, then for all higher orders that vertex remains partitioned in a class by itself. Accordingly, in Fig. 1, only neighborhoods for vertices that partitioned at order n into classes containing more than one vertex are shown at order n+1. Thus, since C_1 and C_4 partition in different classes singly at order 1, they are not shown in

Fig. 1 at order 2 or 3. For the first order neighborhood, the calculation of the information-theoretic indices are shown below:

$$IC_1 = -[1/13 \log_2 1/13 + 2/13 \log_2 2/13 + 1/13 \log_2 1/13 + 1/13 \log_2 1/13 + 7/13 \log_2 7/13 + 1/13 \log_2 1/13] = 2.0349 \text{ bits}$$

$$SIC_1 = IC_1/\log_2 n = 2.0349/\log_2 13 = 0.5499$$

$$CIC_1 = \log_2 n - IC_1 = \log_2 13 - 2.0349 = 1.6655 \text{ bits}$$

BIOLOGICAL APPLICATIONS

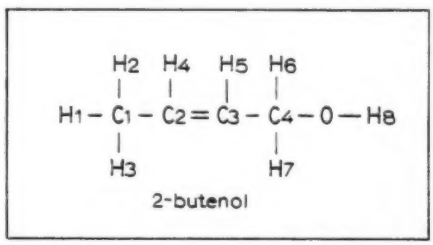
Alcohols

Aliphatic alcohols bring about their characteristic toxic action through the nonspecific perturbation of the biotargets (Roth, 1980). In an earlier study, Basak and Magnuson (1983) found that the aquatic toxicity of alkanols in Pimephales promelas is well correlated with their CIC values. Here we have made a comparative study of the different orders of IC, SIC and CIC indices in correlating the LC_{50} values of ten alcohols, the experimental data being taken from Veith, Call and Brooke (1983). Of the various information indices, CIC_1 and IC_0 gave the best linear correlations with the experimental LC_{50} values.

TABLE 2 Aquatic Toxicity and Information Indices for Aliphatic Alcohols

Name	log LC_{50}	CIC_1	IC_0
Methanol	-0.06	0.79	1.25
Ethanol	-0.51	1.29	1.22
2-Propanol	-0.80	1.80	1.19
1-Butanol	-1.63	2.04	1.16
1-Hexanol	-3.02	2.67	1.12
1-Octanol	-4.00	3.15	1.09
1-Nonanol	-4.40	3.35	1.07
1-Decanol	-4.84	3.52	1.06
1-Undecanol	-5.22	3.68	1.06
1-Dedecanol	-5.27	3.83	1.05

The following equations correlate LC_{50} values with CIC_1 and IC_0, respectively:



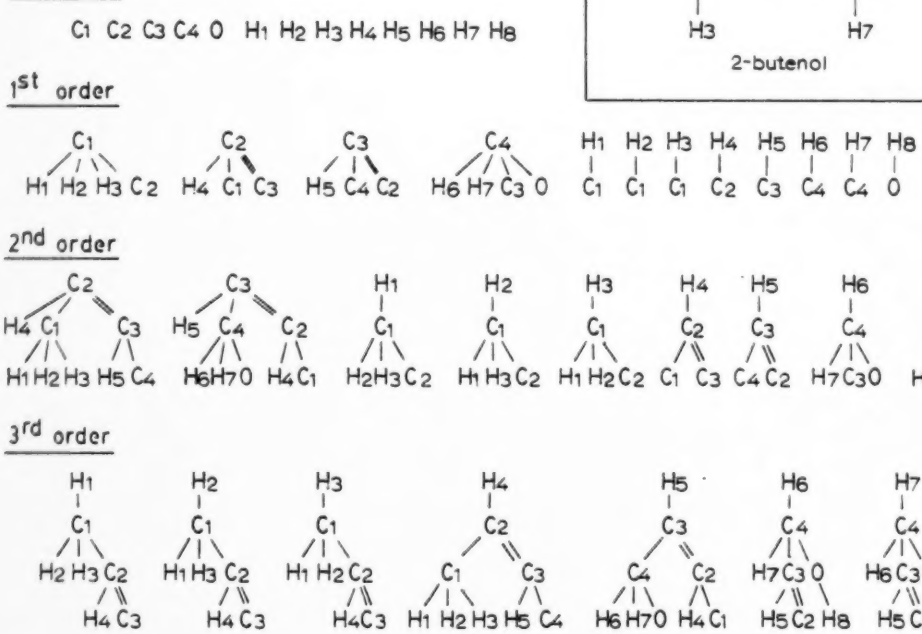


Fig. 1. Labeling and Neighborhoods for 2-Butenol.

	r	s	
$\log LC_{50}$=1.98-1.90 CIC_1	0.99	0.32	(2)
$\log LC_{50}$=-33.8+27.4 IC_0	0.99	0.35	(3)

In the above equations, r is the correlation coefficient and s is the standard deviation.

CONCLUSIONS

It is evident from the results presented above that the measures of neighborhood complexity proposed in this paper reflect important structural aspects which are biologically relevant. In recent years, symmetry has been found to be useful in predicting the physical and biological properties of congeners (Bonchev, Kamenski and Kamenska, 1976; Kier, 1980). The IC, SIC and CIC indices, being quantitative descriptors of neighborhood symmetry, are likely to be useful in biological correlations where symmetry has a vital role in determining bioactivities. Further studies are needed to evaluate the utility of these indices in biomedicinal structure-activity studies.

REFERENCES

Balaban, A. T., A. Chiriac, I. Motoc and Z. Simon (1980). Steric fit in quantitative structure-activity relations. In G. Berthier, M. J. S. Dewar, H. Fischer, K. Fukui, H. Hartmann, H. H. Jaffé, J. Jortner, W. Kutzelnigg, K. Ruedenberg, E. Scrocco, and W. Zeil (Eds.), Lecture Notes in Chemistry, Vol. 15, Springer-Verlag, Berlin-Heidelberg-New York.

Basak, S. C., D. P. Gieschen, V. R. Magnuson and D. K. Harriss (1982). Structure-activity relationships and pharmacokinetics: a comparative study of hydrophobicity, van der Waals' volume and topological parameters. IRCS Med. Sci., 10, 619-620.

Basak, S. C., D. K. Harriss and V. R. Magnuson (1983). A comparative study of lipophilicity versus topological molecular descriptors in biological correlations. J. Pharm. Sci. (in press).

Basak, S. C., and V. R. Magnuson (1983). Molecular topology and narcosis: a quantitative structure-activity relationship (QSAR) study of alcohols using complementary information content (CIC). Arzneim. Forsch. (in press).

Basak, S. C., A. B. Roy and J. J. Ghosh (1980). Study of the structure-function relationship of pharmacological and toxicological agents using information theory. In X. J. R. Avula, R. Bellman, Y. L. Luke, and A. K. Rigler (Eds.), Proceedings of the second international conference on mathematical modeling, Vol. II. University of Missouri-Rolla, Rolla, pp. 851-856.

Bonchev, D., D. Kamenski and V. Kamenska (1976). Symmetry and information content of chemical structures. Bull. Math. Biol., 38, 119-133.

Bonchev, D., O. Mekenyan and N. Trinajstić (1981). Isomer discrimination by topological information approach. J. Comput. Chem., 2, 127-148.

Burnside, W. (1911). Theory of Groups of Finite Order. Dover Publications Inc., New York.

Kier, L. B. (1980). Use of molecular negentropy to encode structure governing biological activity. J. Pharm. Sci., 69, 807-810.

Kier, L. B., and L. H. Hall (1976). Molecular Connectivity in Chemistry and Drug Research. Academic Press, New York - London.

Mowshowitz, A. (1968a). Entropy and the complexity of graphs: I. an index of the relative complexity of a graph. Bull. Math. Biophys., 30, 175-203.

Mowshowitz, A. (1968b). Entropy and the complexity of graphs: II. the information content of digraphs and infinite graphs. Bull. Math. Biophys., 30, 225-240.

Randić, M., and C. L. Wilkins (1979). Graph theoretical ordering of structures as a basis for systematic searches for regularities in molecular data. J. Phys. Chem., 83, 1525-1540.

Rashevsky, N. (1955). Life, information theory and topology. Bull. Math. Biophys., 17, 229-235.

Raychaudhury, C., S. C. Basak, S. K. Ray, A. B. Roy and J. J. Ghosh (1982). Paper presented at the 19th annual meeting of the Society of Engineering Sciences, Inc., University of Missouri-Rolla, Rolla.

Read, R. C., and D. G. Cornell (1977). The graph isomorphism disease. Journal of Graph Theory, 1, 339-363.

Roth, S. H. (1980). Membrane and cellular action of anesthetic agents. Fed. Proc., 39, 1595-1599.

Sarkar, R., A. B. Roy and P. K. Sarkar (1978). Topological information content of genetic molecules. Math. Biosci., 39, 299-312.

Shannon, C. E., and W. Weaver (1949). The Mathematical Theory of Communication. University of Illinois Press, Urbana.

Trucco, E. (1956). A note on the information content of graphs. Bull. Math. Biophys., 18, 129-135.

Veith, G. D., D. J. Call and L. T. Brooke (1983). Structure-toxicity relationship for the fathead minnows: narcotic industrial chemicals. Can. J. Fish., Aq. Sci. (in press).

ACKNOWLEDGMENT

The authors wish to express their appreciation to the U. S. Environmental Protection Agency (Contract CR807566) for support.